arXiv:0805.0828v1 [math.OC] 7 May 2008

Gradient-like observers for invariant dynamics on a Lie group

C. Lageman*, J. Trumpf† and R. Mahony‡

May 7, 2008

Abstract

This paper proposes a design methodology for non-linear state observers for invariant kinematic systems posed on finite dimensional connected Lie groups, and studies the associated fundamental system structure. The concept of synchrony of two dynamical systems is specialised to systems on Lie groups. For invariant systems this leads to a general factorisation theorem of a nonlinear observer into a synchronous (internal model) term and an innovation term. The synchronous term is fully specified by the system model. We propose a design methodology for the innovation term based on gradient-like terms derived from invariant or non-invariant cost functions. The resulting nonlinear observers have strong (almost) global convergence properties and examples are used to demonstrate the relevance of the proposed approach.

Keywords: Observers, Lie groups, Synchrony, Gradient systems.

1 Introduction

There has been a surge of interest recently in nonlinear observer design for systems on Lie groups. A driver for recent work is the growing demand for highly robust state estimation algorithms for autonomous robotic systems such as unmanned aerial, ground or submersible vehicles. Nonlinear observer design for such applications offers the potential of computationally simple state-estimation algorithms with strong robustness and global stability guarantees; as compared to the alternative of nonlinear filter designs (eg. extended Kalman filters [1] or particle filters [2]) that provide more information (*posterior* distributions for

*C. Lageman is with the Centre for Mathematics and its Applications, The Australian National University, Canberra ACT 0200, Australia, `lageman@maths.anu.edu.au`. His work is supported by the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

†J. Trumpf is with the Department of Information Engineering, The Australian National University, Canberra ACT 0200, Australia, `Jochen.Trumpf@anu.edu.au`.

‡R. Mahony is with the Department of Engineering, The Australian National University, Canberra ACT 0200, Australia, `Robert.Mahony@anu.edu.au`.

the state estimates) but usually require significant computational resources and rarely have strong global stability or robustness guarantees. Previous work in nonlinear observer design for systems on Lie groups is often closely linked to specific applications. In the early nineties, Salcudean [3] proposed a nonlinear observer for the attitude estimation of a rigid-body using the unit quaternion representation of the special-orthogonal Lie group $SO(3)$. This work is seminal to a series of papers that develop nonlinear attitude observers for rigid-body dynamics [4, 5, 6, 7, 8, 9, 10], exploiting either the unit quaternion group structure or the rotation matrix Lie-group structure of $SO(3)$. The resulting attitude observers have comparable performance to state-of-the-art nonlinear filtering techniques [11], generally have much stronger global stability and robustness properties, and are simple to implement. The full pose estimation problem has also attracted recent attention [12, 13, 14, 15]. In this case the underlying state space is the special Euclidean group $SE(3)$ comprising both attitude and translation of a rigid-body. Theoretical work in this direction is less advanced due to the more complex algebraic and geometric structure of $SE(3)$. Recently, several authors have made a start in developing a theoretical foundation for observer design for general systems with invariance properties, and in particular systems with a Lie-group state space and the natural left invariant dynamics [16, 17]. Early work in this direction considered uniform invariance properties across the system, measurements and observer design. More recent work recognises that it may be necessary to consider different invariance properties for the system than the measurements in order to obtain well conditioned observers [18]. Results in this area are very recent and there is no well established observer design methodology for invariant systems on a Lie group, even in the case of full state measurement.

In this paper, we study the design of non-linear observers for state space systems where the state is evolving on a finite dimensional, connected Lie group. We consider the case where full measurements are available for the system kinematics and provide a characterisation of invariant observer structures leading to an observer design methodology. Traditional full state observers, for systems evolving on vector spaces, employ a design paradigm that goes back to the work of Kalman [19] and Luenberger [20]: The observer system is designed as a combination of a copy of the system or *internal model* (i.e. a part that can in principle replicate the observed system's trajectory), plus an innovation term which serves to drive the observer trajectory towards the correct system trajectory in the presence of initialization or measurement errors. We build on the results presented in [16, 17] to systematically study the invariance properties of internal models and innovation terms inherent in nonlinear observer design on Lie groups. We define the concept of synchrony of a plant/observer pair of systems (cf. also [21]) with Lie-group state-space in terms of constant error evolution for a canonical invariant error term that is defined. This leads to a definition of internal model and a very general concept of innovation term for plant/observer systems on Lie groups. The internal model structure is shown to be a copy of the invariant plant dynamics. To design the innovation term, we utilise gradient like driving terms derived from algebraic cost functions on

the Lie group. Under mild assumptions, we prove almost global exponential convergence of the resulting observers and provide example derivations for the important applications of attitude and pose observer design for rigid-body kinematics. Finally, we link gradient like dynamics of the canonical error back to a specific structure of the observer consisting of a synchronous term (internal model) and a gradient-like innovation term. Thus, the paper provides a coherent theory of nonlinear observer design for systems with left (resp. right) invariant kinematics on a Lie group for which one has a right (resp. left) invariant Morse-Bott cost function and full measurements.

After the introduction, Section 2 provides an overview of the systems considered and notation used. Section 3 introduces the concept of synchrony, defines left and right invariant errors, and proves some results concerning the canonical nature of invariant errors. Section 4 defines a concept of internal model and innovation term for systems on Lie groups. Section 5 goes on from the characterisation of observers to propose a specific design methodology based on using gradients of cost functions to define the innovation terms in the general observer structure defined in Section 4. Examples of observer designs on $SO(3)$ and $SE(3)$ are provided. The final technical section, Section 6, provides the link from gradient error dynamics back to the structure of the observer, confirming that the design methodology proposed in Section 5 is the only way to obtain the natural gradient error dynamics associated with a known cost function. A short paragraph of conclusions is also provided in Section 7.

A preliminary, short version of this paper has been accepted for publication at the MTNS 2008 [22].

2 Notation and problem formulation

Let G be a finite dimensional, connected Lie group with Lie algebra $\mathfrak{g}$. Denote the identity element in G by e, and left and right multiplication with an element $X \in G$ by L_X and R_X, respectively. The tangent space $T_X G$ of G at X is represented by left or right translations of the Lie algebra, i.e. $T_e L_X \mathfrak{g}$ or $T_e R_X \mathfrak{g}$. We use the simplified notation Xv for vectors $T_e L_X v \in T_X G$ and vX for vectors $T_e R_X v \in T_X G$ with $v \in \mathfrak{g}$. Furthermore, we assume that there is a Riemannian metric $\langle \cdot, \cdot \rangle$ on G. Some of the results presented in the paper will depend on invariance properties of the metric, however, these will be stated explicitly where required and there are no general assumptions made regarding invariance of the metric. We use the norm $||v||^2 = \langle v, v \rangle$ for $v \in \mathfrak{g}$.

Consider a left invariant system on G of the form

$$\dot{X} = Xu, \tag{1}$$

where $u \colon \mathbb{R} \to \mathfrak{g}$ is a function termed the input signal. An input u of system (1) is *admissible* if solutions of the system are unique, exist for all time and are sufficiently smooth. The system (1) has an equivalent right invariant representation

$$\dot{X} = vX \tag{2}$$

with $v\colon \mathbb{R} \to \mathfrak{g}$ an admissible input. The input signals u and v for left and right invariant representations of the same system are related by $v = T_X R_{X^{-1}} T_e L_X u = \operatorname{Ad}_X u$.

This paper discusses the design of observers that fuse potentially noisy measurements of X and u (resp. v) into an estimate of the state X.

Example 1 *Consider the design of an observer to estimate the attitude of a rigid-body [4, 5, 6, 7, 8, 9, 10]. The attitude of the rigid body is an element of the special orthogonal group* $\mathrm{SO}(3)$*, represented by real, orthogonal* 3×3 *matrices. The Lie algebra of* $SO(3)$*, denoted* $\mathfrak{so}(3)$*, is the set of real, skew-symmetric* 3×3 *matrices. The derivative of a curve* $R\colon \mathbb{R} \to \mathrm{SO}(3)$ *coincides with its derivative in* $\mathbb{R}^{3\times 3}$*. The maps* $T_R L_S : \ R\Omega \mapsto SR\Omega$ *and* $T_R R_S : \ R\Omega \mapsto R\Omega S$ *are given by left and right multiplication of the matrices in* $T_R \,\mathrm{SO}(3)$ *with* S*, respectively. The tangent spaces* $T_R \,\mathrm{SO}(3)$ *are identified with*

$$T_R \,\mathrm{SO}(3) \equiv \{R\Omega \mid \Omega \in \mathfrak{so}(3)\} \subset \mathbb{R}^{3\times 3}.$$

The special orthogonal group has a bi-invariant Riemannian metric induced by the Euclidean metric on a the skew symmetric matrices, i.e. $\langle R\Omega, R\Pi\rangle = \mathrm{tr}(\Omega^\top \Pi)$.

Consider the left invariant dynamics

$$\dot{R} = R\Omega$$

on $\mathrm{SO}(3)$ *with* $\Omega\colon \mathbb{R} \to \mathfrak{so}(3)$ *admissible. This system models the kinematics of the attitude* R *of a coordinate frame fixed to a rigid body in 3D-space relative to an inertial frame. Here,* $\Omega \in \mathfrak{so}(3)$ *encodes the angular velocity* $\omega \in \mathbb{R}^3$ *of the rigid-body measured in the body-fixed frame*

$$\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

In robotic applications, measurements of the attitude R *are provided by exteroceptive sensor systems such as magnetometers, accelerometers and vision systems. The measurement of* Ω *is typically obtained from on-board gyrometer systems.*

Example 2 *The above example can easily be extended to model the full pose of a rigid-body, including the position* p *in 3D-space, by considering invariant systems on the special Euclidean group* $\mathrm{SE}(3)$ *[12, 13, 14, 15]. The special Euclidean group* $\mathrm{SE}(3)$ *has a representation as a semidirect product [23] of* $\mathrm{SO}(3)$ *and* $\mathbb{R}^3$

$$\mathrm{SE}(3) = \{(R, p) \mid R \in \mathrm{SO}(3), p \in \mathbb{R}^3\},$$

where the group product is given by

$$(R, p)(S, q) = (RS, p + Rq).$$

Recall the shorthand notation

$$T_e L_{\begin{pmatrix} R & p \end{pmatrix}} \begin{pmatrix} \Omega & V \end{pmatrix} = \begin{pmatrix} R & p \end{pmatrix} \begin{pmatrix} \Omega & V \end{pmatrix}$$

with $\begin{pmatrix} R & p \end{pmatrix} \in \mathrm{SE}(3)$ *and* $\begin{pmatrix} \Omega & V \end{pmatrix} \in \mathfrak{se}(3)$ *and consider the left invariant system*

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{pmatrix} R(t) & p(t) \end{pmatrix} = \begin{pmatrix} R & p \end{pmatrix} \begin{pmatrix} \Omega & V \end{pmatrix} = \begin{pmatrix} R\Omega & RV \end{pmatrix}$$

with $\Omega \colon \mathbb{R} \to \mathfrak{so}(3)$ *and* $V : \mathbb{R} \to \mathbb{R}^3$ *admissible. Here,* V *is the linear velocity of the rigid-body measured in the body-fixed frame.*

In robotic applications, the measurement of p *is provided by exteroceptive sensor systems such as GPS (global positioning systems), radar or vision systems. The measurement of linear velocity* V *is more challenging as there are few sensor systems that directly measure linear velocity. Typical measurement systems depend on fusing differentiated exteroceptive position measurements with integration of accelerometer measurements.*

The two examples (Ex. 1 and 2) provide an excellent means to demonstrate the structure of the design principles proposed in this paper and indicate the relevance of the results to real-world engineering problems. It is not the purpose of this paper, however, to enter into the details of the practical implementation of the proposed observers for a real world system and we will not consider issues such as choice of sensor systems or characterisation of noise any further than the brief discussion above and a couple of remarks later in the paper.

3 Synchrony and error functions

In this section, we introduce the general concept of *synchrony* between pairs of systems that evolve on a given Lie group G and have a common input. Loosely, synchrony refers to an equivalence, but not equality, of trajectories of two dynamical systems. More formally, we will define synchrony of two systems in terms of an *error function* E that quantifies the instantaneous difference between the system trajectories. The pair of systems is called E*-synchronous* if the error E is constant along trajectories. We show that synchronous invariant systems on Lie-groups have a particular structure and that the error function E associated with invariant synchronous systems is also structurally constrained.

Consider a pair of systems on G driven by the same admissible input u

$$\dot{X} = F_X(X, u, t), \tag{3}$$

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, u, t) \tag{4}$$

with $F_X, F_{\hat{X}} : G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_X(X, u, t) \in T_X G$ and $F_{\hat{X}}(\hat{X}, u, t) \in T_{\hat{X}} G$. The solutions of (3) and (4) are denoted by $X(t; X_0, u)$ and $\hat{X}(t; \hat{X}_0, u)$, respectively.

Definition 3 *Consider the systems* (3) *and* (4). *Let* $E\colon G \times G \to M$ *be a smooth* error function*,* M *a smooth manifold. The systems* (3) *and* (4) *are said to be* E-synchronous *if for all admissible* $u\colon \mathbb{R} \to \mathfrak{g}$*, all initial values* $X_0, \hat{X}_0 \in G$ *of* (3) *and* (4) *and all* $t \in \mathbb{R}$

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\hat{X}(t; \hat{X}_0, u), X(t; X_0, u)) = 0.$$

Two particularly simple error functions on a Lie group G are the canonical *right invariant* error

$$E_r(\hat{X}, X) := \hat{X} X^{-1} \tag{5}$$

and the canonical *left invariant* error

$$E_l(\hat{X}, X) := X^{-1} \hat{X}. \tag{6}$$

Where the arguments $\hat{X}$ and X are clear from the context we simply write E_r and E_l. The label "invariant" refers to simultaneous state space transformations of both systems. That is, for all $X, \hat{X}, S \in G$, one has $E_r(\hat{X}S, XS) = E_r(\hat{X}, X)$ and an analogous result for E_l.

Remark 4 *Observe that both* E_r *and* E_l *are non-degenerate in the sense that the partial maps* $E(\hat{X}, \cdot)\colon G \to M$ *and* $E(\cdot, X)\colon G \to M$ *are global diffeomorphisms (here,* $M = G$*).*

The next proposition shows that any error E for which two *invariant* systems are E-synchronous, can be factored into one of the canonical invariant errors concatenated with a map from the Lie group into a manifold. Thus, the right and left invariant errors can be thought of as the fundamental error functions for pairs of invariant systems.

Proposition 5 *Consider the pair of invariant systems on* G

$$\dot{X} = Xu, \tag{7}$$

$$\dot{\hat{X}} = \hat{X} u \tag{8}$$

for a single admissible input $u\colon \mathbb{R} \to \mathfrak{g}$*. Let* M *be a smooth manifold. If there exists a smooth error function* $E\colon G \times G \to M$ *such that the systems* (7) *and* (8) *are* E*-synchronous, then* E *has the form*

$$E(\hat{X}, X) = g(\hat{X} X^{-1}) = g(E_r(\hat{X}, X)),$$

where $g\colon G \to M$ *is a smooth function.*

The analogous result holds for right-invariant systems

$$\dot{X} = vX$$

$$\dot{\hat{X}} = v\hat{X}.$$

In that case one has

$$E(\hat{X}, X) = g(X^{-1} \hat{X}) = g(E_l(\hat{X}, X))$$

Proof: Consider the pair of systems (7) and (8) and an error function E as above. Let $X_0, \hat{X}_0, S \in G$ be arbitrary. Choose a smooth, bounded curve $T\colon \mathbb{R} \to G$ with bounded derivatives such that $T(0) = e$ and $T(1) = S$. Let $(X, \hat{X})$ denote the solution of the pair of systems (7) and (8) for $u(t) = T^{-1}(t)\dot{T}(t)$ and $(X, \hat{X})(0) = (X_0, \hat{X}_0)$. Note that $(X, \hat{X})(t) = (X_0 T(t), \hat{X}_0 T(t))$. The error function E is constant on the trajectory $(X, \hat{X})(t)$ and therefore $E(\hat{X}_0, X_0) = E(\hat{X}_0 S, X_0 S)$. Since $X_0, \hat{X}_0$ and S are arbitrary, the function E is right invariant under the action of G, i.e.

$$E = E \circ R_S \text{ for all } S \in G.$$

But this implies that $E(\hat{X}, X) = E \circ R_{X^{-1}}(\hat{X}, X) = E(\hat{X}X^{-1}, e)$ for all $X, \hat{X} \in G$. Hence E has the form $E(\hat{X}, X) = g(\hat{X}X^{-1})$ where $g\colon G \to M$ is the smooth function $g(Z) = E(Z, e)$. The other case is proven by an analogous argument. $\square$

Given that the invariant errors E_l and E_r play such a key role in the analysis of invariant synchronous systems it is of interest to specialise the notion of synchrony introduced above to the canonical errors.

Definition 6 *A pair of systems* (3) *and* (4) *is termed* right synchronous *if they are E_r-synchronous and* left synchronous *if they are E_l-synchronous.*

In the remainder of the section, we consider the structure of a pair of synchronous systems. We are interested in the case where the "first" system is left (resp. right) invariant (1) (resp. (2)). We consider a "partner" system that is either left or right synchronous and derive constraints on the structure of the partner system. The result will provide us with the first half of the template for the design of non-linear observers. The proof of the main result (Theorem 8) uses the following lemma.

Lemma 7 *Let $X\colon \mathbb{R} \to G$ and $Y\colon \mathbb{R} \to G$ be two smooth curves. Then*

$$\begin{aligned}
(\dot{X^{-1}})Y &= -T_{X^{-1}} R_Y T_e L_{X^{-1}} T_X R_{X^{-1}} \dot{X} + T_Y L_{X^{-1}} \dot{Y}, \\
\dot{Y} X^{-1} &= T_Y R_{X^{-1}} \dot{Y} - T_{X^{-1}} L_Y T_e L_{X^{-1}} T_X R_{X^{-1}} \dot{X}.
\end{aligned}$$

Proof: Recall that for the inverse function $\mathrm{inv}(X) = X^{-1}$ we have $T_X\,\mathrm{inv} = -T_X\,(L_{X^{-1}} R_{X^{-1}})$. The equations for the derivatives now follow from the usual calculus rules for the multiplication on Lie groups. $\square$

Theorem 8 *Consider the left invariant system* (1) *and let a second system be given by the general expression*

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, u, t). \tag{9}$$

The systems (1) *and* (9) *are right synchronous if and only if*

$$F_{\hat{X}}(\hat{X}, u, t) = \hat{X}u. \tag{10}$$

The systems are left synchronous if and only if

$$F_{\hat{X}}(\hat{X}, u, t) = \hat{X} \operatorname{Ad}_{\hat{X}^{-1}X} u = \hat{X} \operatorname{Ad}_{E_l^{-1}} u. \tag{11}$$

Proof: By Lemma 7 we have for trajectories $X(t)$ of (1) and $\hat{X}(t)$ of (9) that

$$\begin{aligned}\dot{E}_r &= T_{\hat{X}} R_{X^{-1}} \dot{\hat{X}} - T_{X^{-1}} L_{\hat{X}} T_e L_{X^{-1}} T_X R_{X^{-1}} \dot{X} \\ &= T_{\hat{X}} R_{X^{-1}} F_{\hat{X}}(\hat{X}, u, t) \\ &\quad - T_{X^{-1}} L_{\hat{X}} T_e L_{X^{-1}} T_X R_{X^{-1}} T_e L_X u \\ &= T_{\hat{X}} R_{X^{-1}} F_{\hat{X}}(\hat{X}, u, t) - T_{X^{-1}} L_{\hat{X}} T_e R_{X^{-1}} u\end{aligned}$$

and

$$\begin{aligned}\dot{E}_l &= -T_{X^{-1}} R_{\hat{X}} T_e L_{X^{-1}} T_X R_{X^{-1}} \dot{X} + T_{\hat{X}} L_{X^{-1}} \dot{\hat{X}} \\ &= -T_{X^{-1}} R_{\hat{X}} T_e L_{X^{-1}} T_X R_{X^{-1}} T_e L_X u \\ &\quad + T_{\hat{X}} L_{X^{-1}} F_{\hat{X}}(\hat{X}, u, t) \\ &= -T_{X^{-1}} R_{\hat{X}} T_e R_{X^{-1}} u + T_{\hat{X}} L_{X^{-1}} F_{\hat{X}}(\hat{X}, u, t).\end{aligned}$$

Hence E_r is constant if and only if

$$F_{\hat{X}}(\hat{X}, u, t) = T_e L_{\hat{X}} u = \hat{X}u.$$

and E_l is constant if and only if

$$\begin{aligned}F_{\hat{X}}(\hat{X}, u, t) &= T_{X^{-1}\hat{X}} L_X T_{X^{-1}} R_{\hat{X}} T_e R_{X^{-1}} u \\ &= T_e L_{\hat{X}} \operatorname{Ad}_{\hat{X}^{-1}X} u = \hat{X} \operatorname{Ad}_{E_l^{-1}} u.\end{aligned}$$

□

In the case where the "first" system is right invariant the synchronous terms for the different errors are interchanged.

Theorem 9 *Consider the right invariant system* (2) *and let a second system be given by*

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, v, t). \tag{12}$$

The systems (2) *and* (12) *are left synchronous if and only if*

$$F_{\hat{X}}(\hat{X}, v, t) = v\hat{X}.$$

The systems are right synchronous if and only if

$$F_{\hat{X}}(\hat{X}, v, t) = (\operatorname{Ad}_{\hat{X}X^{-1}} v)\hat{X} = (\operatorname{Ad}_{E_r} v)\hat{X}.$$

It is interesting to observe that for a left invariant system (1) and choosing the right synchronous constraint, one obtains left invariant dynamics for the partner system (10). The dynamics (10) are highly desirable as the fundamental component of a non-linear observer design since they are independent of the observed system's state X and yet ensure that the observer dynamics "track" the system state. This contrasts to choosing the left synchronous constraint, where the resulting partner system structure (11) depends on the error $E_l = X^{-1}\hat{X}$ that requires knowledge of the observed system's state X and cannot be implemented in a real observer system. Thus, it is natural to consider right invariant errors and right synchronicity when designing observers for left invariant systems (1); and vice-versa for right invariant systems (2).

4 Internal models and innovation terms

In this section, we define the concept of internal model specialised for Lie group systems. This definition is used to show that any observer of an invariant system on a Lie group containing an internal model, can be split into a synchronous term, providing the internal model properties of the observer, and a second term that we identify as an innovation term. Thus, a natural approach to observer design is to choose the observer as a sum of a synchronous (internal model) term plus an innovation term, analogous to the "internal model plus innovation term" design paradigm for linear systems.

Consider an observed system

$$\dot{X} = F_X(X, u, t) \tag{13}$$

with $F_X\colon G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_X(X, u, t) \in T_X G$ and an observer

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, Y, w, t) \tag{14}$$

with $F_{\hat{X}}\colon G \times G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_{\hat{X}}(\hat{X}, Y, w, t) \in T_{\hat{X}}G$. Note that the observer (14) has two inputs Y and w, with Y to be fed with measurements of X and w to be fed with measurements of u, respectively. The idea behind an internal model is that the observer should be able to replicate the exact trajectory of the observed system if it is provided with exact input information; that is, the actual initial condition of the observed system's state and exact measurements. There is no a priori requirement that the system and the observer are identical dynamical systems, only that they correspond along certain very specific trajectories.

Definition 10 *Consider the pair of systems* (13) *and* (14). *One says that* (14) *has an* internal model *of* (13) *if for all admissible* $u\colon \mathbb{R} \to G$, $X_0 \in G$ *and all* $t \in \mathbb{R}$

$$\hat{X}(t; X_0, X(t; X_0, u), u) = X(t; X_0, u), \tag{15}$$

where $X(t; X_0, u)$ *and* $\hat{X}(t; \hat{X}_0, Y, w)$ *denote the solutions of* (13) *and* (14), *respectively.*

Note that by Theorem 8 a right synchronous system for the left invariant system (1) has the form

$$\dot{\hat{X}} = \hat{X}w \tag{16}$$

with $w = u$. It is straightforward to verify that (16) has an internal model of the original left invariant system. The following result shows that any observer of a left invariant system containing an internal model can be split into two parts; the first or which is right synchronous with the original system, providing the internal model properties of the observer, and a second term that will be identified as an innovation term.

Theorem 11 *Consider the left invariant system* (1) *and the observer* (14)*. Assume that the observer has an internal model of the system. Then the right hand side of the observer can be written*

$$F_{\hat{X}}(\hat{X}, Y, w, t) = \hat{X}w + \alpha(\hat{X}, Y, w, t), \tag{17}$$

where $\alpha : G \times G \times \mathfrak{g} \times \mathbb{R} \to TG$ *is a smooth function satisfying* $\alpha(\hat{X}, Y, w, t) \in T_{\hat{X}}G$ *and*

$$\alpha\left(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t\right) = 0 \tag{18}$$

for all admissible $u \colon \mathbb{R} \to \mathfrak{g}$*,* $X_0 \in G$ *and* $t \in \mathbb{R}$*.*

Proof: Define $\alpha(\hat{X}, Y, w, t) = F_{\hat{X}}(\hat{X}, Y, w, t) - \hat{X}w$ for all $\hat{X}, Y \in G$, $w \in \mathfrak{g}$ and $t \in \mathbb{R}$. Differentiating the internal model equation (15) and using the system and observer equations yields

$$F_{\hat{X}}\left(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t\right) = \\ X(t; X_0, u)u$$

for all admissible $u \colon \mathbb{R} \to \mathfrak{g}$, $X_0 \in G$ and $t \in \mathbb{R}$. We hence get

$$\begin{aligned} \alpha\left(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t\right) &= \\ F_{\hat{X}}\left(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t\right) &- \\ \hat{X}(t; X_0, X(t; X_0, u), u)u &= \\ X(t; X_0, u)u - \hat{X}(t; X_0, X(t; X_0, u), u)u &= 0 \end{aligned}$$

where the last equality follows again from the definition of an internal model. $\square$

A similar result can obviously be obtained for right invariant observed systems.

The decomposition (17) is analogous to the internal model plus innovation term decomposition of linear observers. The concept of right (resp. left) synchronicity provides a structural characterisation of internal models for left

(resp. right) invariant systems. The α term in Equation (17) is analogous to the innovation term in classical linear observer design. Condition (18) means that α is zero along corresponding trajectories of the system and the observer.

Definition 12 *Consider the pair of systems* (13) *and* (14) *and assume that* (14) *has an internal model of* (13). *We call a map* $\alpha\colon G \times G \times \mathfrak{g} \times \mathbb{R} \to TG$ *an* innovation term *if*

(C1) $\alpha(\hat{X}, Y, w, t) \in T_{\hat{X}}G$ *for all* $\hat{X}, Y \in G$*,* $w \in \mathfrak{g}$*,* $t \in \mathbb{R}$ *and*

(C2) $\alpha(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t) = 0$ *for all admissible* $u\colon \mathbb{R} \to \mathfrak{g}$*,* $X_0 \in G$ *and* $t \in \mathbb{R}$*.*

Note that the conditions specified in Definition 12 are the least restrictive possible. In general, an innovation term must be chosen carefully to ensure that the trajectory of the observed system is an asymptotically stable limit set of the observer trajectory. We discuss design of the innovation term in Section 5.

Note further that Condition (C2) in Definition 12 is in particular implied by the following stronger condition

(C2') $\alpha(Y, Y, w, t) = 0$ for all $Y \in G$, $w \in \mathfrak{g}$ and $t \in \mathbb{R}$.

Conditions (C2) and (C2') are in general not equivalent since there is no guarantee that for a given admissible input u the trajectories of the system (and/or those of the observer) will cover all of G at any given time.

In summary, we propose the following structure for the design of non-linear observers for invariant systems on Lie groups.

$$\dot{\hat{X}} = \hat{X} w_l + \alpha(\hat{X}, Y, w_l, t) \quad \text{(left observer)} \tag{19}$$

This observer is intended for left invariant observed systems of the form (1), with Y receiving measurements of X and w_l receiving measurements of u. Note that this observer is in general not left invariant, since α need not be left invariant.

$$\dot{\hat{X}} = w_r \hat{X} + \alpha(\hat{X}, Y, w_r, t) \quad \text{(right observer)} \tag{20}$$

Correspondingly, this observer is intended for right invariant observed systems of the form (2), with Y receiving measurements of X and w_r receiving measurements of v. Note that this observer is in general not right invariant.

5 Gradient observers

In this section, we present an approach to choosing innovation terms for non-linear observers of the form (19) and (20) based on using the gradient descent direction of a suitable cost function.

Let $f\colon G \times G \to \mathbb{R}$ be a smooth, non-negative cost function. Furthermore, let the diagonal $\Delta = \{(X, X) \mid X \in G\}$ consist of global minima of f. Recall

that the Riemannian gradient of f with respect to the product metric $\langle\cdot,\cdot\rangle_p$ on $G \times G$ is defined by

$$\langle \operatorname{grad} f(\hat{X}, Y), (\eta, \zeta)\rangle_p = df(\hat{X}, Y)(\eta, \zeta)$$

for all $\hat{X}, Y \in G$, $\eta \in T_{\hat{X}}G$, $\zeta \in T_Y G$. Since we use the product metric, the gradient splits into the gradients with respect to the first and second parameter, i.e.

$$\langle \operatorname{grad} f(\hat{X}, Y), (\eta, \zeta)\rangle_p = \langle \operatorname{grad}_1 f(\hat{X}, Y), \eta\rangle + \langle \operatorname{grad}_2 f(\hat{X}, Y), \zeta\rangle.$$

We propose to use the gradient of f with respect to the first parameter as an innovation term α in the design of observers. Given the observed system (1) (resp. (2)) then the design of the observer is

$$\dot{\hat{X}} = \hat{X} w_l - \operatorname{grad}_1 f(\hat{X}, Y), \quad \text{(left observer)} \tag{21}$$

respectively

$$\dot{\hat{X}} = w_r \hat{X} - \operatorname{grad}_1 f(\hat{X}, Y). \quad \text{(right observer)} \tag{22}$$

It remains to analyze the dynamics associated with this choice and show that the observer trajectory asymptotically converges to the observed system's trajectory.

5.1 Error dynamics

To understand the stability of the proposed observers (21) and (22) we analyze the case where exact measurements of the input u (resp. v) of system (1) (resp. (2)) as well as exact measurements of the state X are available. In terms of the variables introduced in Section 4 one has $Y = X$ and $w_l = u$ (resp. $w_r = v$) in the case of the observer for a left (resp. right) invariant system.

The following result is used in the development later in the section.

Lemma 13 *Let G be a Lie group. Let $f : G \times G \to \mathbb{R}$ be a left invariant cost function and take a left invariant Riemmanian metric. Then for all $X, Y, Z \in G$*

$$T_X L_{Y^{-1}} \operatorname{grad}_1 f(X, YZ) = \operatorname{grad}_1 f(Y^{-1}X, Z)$$

If f and the Riemmanian metric are right invariant, then for all $X, Y, Z \in G$

$$T_X R_{Z^{-1}} \operatorname{grad}_1 f(X, YZ) = \operatorname{grad}_1 f(XZ^{-1}, Y)$$

Proof: If f is left invariant, then $f \circ L_Y = f$. Using this fact and the standard rules for transformations of Riemannian gradients, we have that

$$\begin{aligned}\operatorname{grad}_1 f(Y^{-1}X, Z) &= \operatorname{grad}_1 (f \circ L_Y)(Y^{-1}X, Z) \\ &= (T_{Y^{-1}X} L_Y)^* \operatorname{grad}_1 f(X, YZ),\end{aligned}$$

where $(T_{Y^{-1}X} L_Y)^*$ denotes the Hilbert space adjoint of the linear map $T_{Y^{-1}X} L_Y$. Since the Riemannian metric is left invariant, we have that $(T_{Y^{-1}X} L_Y)^* = T_X L_{Y^{-1}}$. The right invariant case follows from an analogous argument. $\square$

To analyze the asymptotic stability of the observer trajectory to the observed system's trajectory it is convenient to consider the dynamics of the error functions E_r and E_l and prove that their trajectories converge to the identity element of the Lie group. Since E_r and E_l are non-degenerate, this will imply the desired asymptotic stability. In particular, under suitable invariance conditions on the cost function and the Riemannian metric the error dynamics for the "correct" invariant error are autonomous.

Theorem 14 *Consider the left invariant system* (1). *Let $f : G \times G \to \mathbb{R}$ be a right invariant cost function and take a right invariant Riemannian metric on G. Consider the left observer dynamics* (21). *Then the error dynamics of the right invariant error E_r* (5) *are given by*

$$\dot{E}_r = -\operatorname{grad}_1 f(E_r, e).$$

If the right invariant system (2) *is considered with a left invariant cost function and Riemannian metric along with the right observer dynamics* (22), *then the error dynamics of the left invariant error E_l* (6) *are given by*

$$\dot{E}_l = -\operatorname{grad}_1 f(E_l, e).$$

Proof: Directly from the system equations and the equations of the left observer it follows that

$$\begin{aligned}\dot{E}_r &= -T_{\hat{X}} R_{X^{-1}} \operatorname{grad}_1 f(\hat{X}, Y) \\ &= -T_{\hat{X}} R_{X^{-1}} \operatorname{grad}_1 f(\hat{X}, X).\end{aligned}$$

Using Lemma 13 and the right invariance of f and the Riemannian metric, we get that

$$\dot{E}_r = -\operatorname{grad}_1 f(\hat{X} X^{-1}, e) = -\operatorname{grad}_1 f(E_r, e).$$

The result for the right observer follows from an analogous argument. $\square$

The gradient dynamics of the error yield the following convergence result in the noise-free case.

Theorem 15 *Assume that $Y \mapsto f(Y, e)$ is a Morse-Bott function with a global minimum at e and no other local minima. If f and the Riemannian metric on G are both right invariant, then both errors E_r and E_l of the left observer* (21) *converge to e for generic initial conditions. Furthermore, $f(E_r, e)$ converges monotonically to $f(e, e)$ and this convergence is locally exponential near e. If f and the Riemannian metric on G are both left invariant, then both errors E_r and E_l of the right observer* (22) *converge to e for generic initial conditions. Furthermore, $f(E_l, e)$ converges monotonically to $f(e, e)$ and this convergence is locally exponential near e.*

Proof: Let us consider the left observer. The gradient structure of the error dynamics (Theorem 14) and the conditions on the cost function yield the monotonic convergence of $f(E_r, e)$ and therefore the convergence of E_r to e. As f is a Morse-Bott function, the convergence of $f(E_r, e)$ is exponential near e. Since convergence of E_r to e implies the convergence of E_l to e, we get the convergence of the left invariant error, too. The result for the right observer is proved by an analogous argument. $\square$

Remark 16 *Consider the left invariant system* (1) *and noisy measurements* $w_l = u+\delta$ *and* $Y = N_l X$*, with additive driving noise* $\delta \in \mathfrak{g}$ *and left multiplicative state noise* $N_l \in G$*. A straightforward calculation yields*

$$\dot{E}_r = \mathrm{Ad}_{\hat{X}}\, \delta E_r - \mathrm{grad}_1\, f(E_r, N_l)$$

for the canonical right invariant error of the left observer (21)*. There is an analogous formula for the canonical left invariant error of the right observer* (22)*, this time best expressed in terms of right multiplicative state noise* $Y = XN_r$*. It is intuitively clear that suitably bounded noise will yield at least a practical stability result in these cases.*

Theorems 14 and 15 are the main results of this section. Together, these results provide a template for the design of non-linear observers for invariant systems on a Lie-group, Equations (21) and (22). There is still an outstanding question of how to find suitable cost functions that we will address in Section 5.3. In the remainder of this section, we consider some of the special cases that were not addressed in Theorems 14 and 15.

The convergence result in Theorem 15 raises the question of what the dynamics of the other error, i.e. E_l for the left observer and E_r for the right observer, looks like and if its cost converges monotonically to $f(e, e)$. Under suitable invariance conditions, one obtains the following result on the dynamics of the other errors for each observer.

Theorem 17 *Consider the left invariant system* (1)*. Let* $f : G \times G \to \mathbb{R}$ *be a left invariant cost function and take a left invariant Riemannian metric on* G*. Consider the left observer dynamics* (21)*. Then the error dynamics of the left invariant error* E_l *are*

$$\begin{aligned}\dot{E}_l &= (T_e L_{E_l} - T_e R_{E_l})u - \mathrm{grad}_1\, f(E_l, e)\\ &= (E_l u - u E_l) - \mathrm{grad}_1\, f(E_l, e).\end{aligned}$$

If the right invariant system (2) *is considered with a right invariant cost function and Riemannian metric along with the right observer dynamics* (22)*, then the error dynamics of the right invariant error* E_r *are*

$$\begin{aligned}\dot{E}_r &= (T_e R_{E_r} - T_e L_{E_r})v - \mathrm{grad}_1\, f(E_r, e)\\ &= (vE_r - E_r v) - \mathrm{grad}_1\, f(E_r, e).\end{aligned}$$

Proof: Let us consider the left observer. Using Lemmas 7 and 13 one obtains

$$\begin{aligned}
\dot{E}_l &= -T_{X^{-1}} R_{\hat{X}} T_e L_{X^{-1}} T_X R_{X^{-1}} T_e L_X u + T_{\hat{X}} L_{X^{-1}} T_e L_{\hat{X}} u \\
&\quad - T_{\hat{X}} L_{X^{-1}} \operatorname{grad}_1 f(\hat{X}, Y) \\
&= (T_{\hat{X}} L_{X^{-1}} T_e L_{\hat{X}} - T_{X^{-1}} R_{\hat{X}} T_e R_{X^{-1}}) u \\
&\quad - T_{\hat{X}} L_{X^{-1}} \operatorname{grad}_1 f(\hat{X}, X) \\
&= (T_e L_{E_l} - T_e R_{E_l}) u - \operatorname{grad}_1 f(E_l, e).
\end{aligned}$$

The statement for the right observer follows from an analogous calculation. □

It is interesting to consider the case of a bi-invariant cost function f. In this case it is possible to show that the non-gradient term in the error dynamics (Theorem 17) is *passive* with respect to the cost, i.e. an element of the kernel of the differential of the cost.

Lemma 18 *Consider a Lie-group G with Lie-algebra $\mathfrak{g}$. Assume that there exists a bi-invariant cost $f : G \times G \to \mathbb{R}$. Then for all $u, v \in \mathfrak{g}$ and any $E_l, E_r \in G$*

$$\begin{aligned}
\langle (T_e L_{E_l} - T_e R_{E_l}) u, \operatorname{grad}_1 f(E_l, e) \rangle = 0 \\
\langle (T_e R_{E_r} - T_e L_{E_r}) v, \operatorname{grad}_1 f(E_r, e) \rangle = 0
\end{aligned}$$

Proof: For all $u \in \mathfrak{g}$ we have by the bi-invariance of f and standard transformation rules for the gradient that

$$\begin{aligned}
\langle T_e L_{E_l} u, \operatorname{grad}_1 f(E_l, e) \rangle &= \langle u, (T_e L_{E_l})^* \operatorname{grad}_1 f(E_l, e) \rangle \\
&= \langle u, \operatorname{grad}_1 f(e, E_l^{-1}) \rangle \\
&= \langle u, (T_e R_{E_l})^* \operatorname{grad}_1 f(E_l, e) \rangle \\
&= \langle T_e R_{E_l} u, \operatorname{grad}_1 f(E_l, e) \rangle.
\end{aligned}$$

This yields the first equation. The other equation follows from an analogous argument. □

Proposition 19 *Consider the left invariant system* (1)*. Let $f : G \times G \to \mathbb{R}$ be a bi-invariant cost function and take a left invariant Riemannian metric on G. Consider the left observer dynamics* (21)*. Then the time derivative of the cost function along trajectories of* (1) *and* (21) *is given by*

$$\frac{\mathrm{d}}{\mathrm{d}t} f = -|| \operatorname{grad}_1 f(E_l, e) ||^2.$$

The analogous result holds for a right invariant Riemannian metric on G and the right observer dynamics (22)*.*

Proof: Consider the left invariant case. The differential of f is given by

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} f &= \langle \operatorname{grad}_1 f(E_l, e), (T_e L_{E_l} - T_e R_{E_l})u - \operatorname{grad}_1 f(E_l, e)\rangle \\ &= -\langle \operatorname{grad}_1 f(E_l, e), \operatorname{grad}_1 f(E_l, e)\rangle \\ &= -|| \operatorname{grad}_1 f(E_l, e)||^2\end{aligned}$$

by applying Lemma 18. The analogous result for right invariant dynamics is similarly direct. □

The above result can be developed into a convergence result analogous to Theorem 15. In practice, the case that is of most interest is when both the cost function and the metric are bi-invariant, a situation that occurs for the attitude estimation example on $SO(3)$. In this case the following convergence result is obtained.

Corollary 20 *Consider the left invariant system* (1) *and the right invariant system* (2)*. Assume that G admits a bi-invariant metric and let $f : G \times G \to \mathbb{R}$ be a bi-invariant cost function. Assume that $Y \mapsto f(Y, e)$ is a Morse-Bott function with a global minimum at e and no other local minima.*

For both the left and right observers (21) *and* (22) *then $f(E_r, e)$ and $f(E_l, e)$ converge monotonically to $f(e, e)$ for generic initial conditions. The convergence is locally exponential near e.*

Proof: The convergence of the cost of one type error for each filter is shown in Corollary 15. The convergence of the other error follows from the structure of the error dynamics, cf. Theorem 17, and Proposition 19 by a straightforward Lyapunov argument. □

5.2 Example: Attitude estimation on SO(3)

We revisit Example 1 on the special orthogonal group SO(3). The left invariant attitude kinematics are

$$\dot{R} = R\Omega,$$

where R denotes the attitude of a coordinate frame fixed to a rigid body in 3D-space relative to an inertial frame and Ω encodes the angular velocity measured in the body-fixed frame. The velocity measurements are given by $w_l = \Omega$ and the state measurements are given by $Y = R$. The right invariant and left invariant errors have the form $E_r = \hat{R}R^\top$ and $E_l = R^\top \hat{R}$, respectively.

We define the cost function $f(\hat{R}, Y) = \frac{k}{2}\|\hat{R} - Y\|_F^2$, with $\|\cdot\|_F$ the Frobenius matrix norm and k a positive constant. Since the Frobenius norm is invariant under orthogonal transformations it follows that f is bi-invariant. Moreover, the standard Riemannian metric on $SO(3)$, induced by the Euclidean inner product on $\mathbb{R}^{3\times 3}$ restricted to the Lie-algebra of skew symmetric matrices, is also bi-invariant. Let us recall the well-known methods to calculate the gradient of

$\hat{R} \mapsto f(\hat{R}, Y)$, see e.g. [24]. As we use the induced metric on SO(3), the gradient is given by the orthogonal projection of the Euclidean gradient in $\mathbb{R}^{3\times 3}$ to the tangent space of SO(3). This projection to $T_{\hat{R}}SO(3)$ can be readily calculated to be $Z \mapsto \hat{R}\mathbb{P}(\hat{R}^\top Z)$, where $\mathbb{P}$ denotes the orthogonal projection onto the skew-symmetric matrices, that is $\mathbb{P}(Z) = \frac{1}{2}(Z - Z^\top)$. Since the Euclidean gradient of $\hat{R} \mapsto f(\hat{R}, Y)$ is just $k(\hat{R} - Y)$, this yields that $\operatorname{grad}_1 f(\hat{R}, Y) = k\hat{R}\mathbb{P}(I - \hat{R}^\top Y) = -k\hat{R}\mathbb{P}(\hat{R}^\top Y)$. Hence we have the observer

$$\dot{\hat{R}} = \hat{R}w_l + k\hat{R}\mathbb{P}(\hat{R}^\top Y) \quad \text{(left observer)}$$

which coincides with the passive filter

$$\dot{\hat{R}} = \hat{R}\Omega + k\hat{R}\mathbb{P}(\hat{R}^\top R)$$

proposed in [8, 6] (recall that $w_l = \Omega$ and $Y = R$ in the noise free case).

Starting with right invariant attitude kinematics

$$\dot{R} = \Gamma R$$

and measurements $w_r = \Gamma$ and $Y = R$, we have the observer

$$\dot{\hat{R}} = w_r\hat{R} + k\hat{R}\mathbb{P}(\hat{R}^\top Y) \quad \text{(right observer)}$$

which, using $\Gamma = \operatorname{Ad}_R \Omega$, coincides with the direct filter

$$\dot{\hat{R}} = \operatorname{Ad}_R \Omega\hat{R} + k \operatorname{Ad}_{\hat{R}} \mathbb{P}(\hat{R}^\top R)\hat{R}$$

discussed in [8, 6] but also extensively studied over the last ten years by a range of authors [4, 5, 7, 9, 10].

5.3 Construction of invariant cost functions

In this section we investigate the question of finding invariant cost functions on Lie groups.

For a left invariant system, Theorem 14 shows that the canonical right invariant error for the left observer has gradient dynamics if the cost and Riemannian metric are right invariant. A right invariant Riemannian metric can be easily constructed on any Lie group by transporting a scalar product on the Lie algebra to other tangent spaces by right translation. However, it is a priori not clear how to obtain a right invariant cost function, in particular if the group is non-compact. For example, the most natural cost function on $SE(3)$ (cf. Example 2) would be

$$f((R,p),(Y,y)) = \|R - Y\|^2 + \|p - y\|^2, \tag{23}$$

however, it is easily verified that this cost function is not right invariant. It should be noted that finding a Morse-Bott cost function is usually fairly straightforward, (23) is an example of such a function, the challenge lies in ensuring

the function chosen has the desired invariance properties. The next proposition provides a method to obtain an invariant cost function on a Lie group given that a suitable Morse-Bott function on G is available.

Proposition 21 *Let G be a Lie group and let $g\colon G \to \mathbb{R}$ be a smooth function. Then $f\colon G \times G \to \mathbb{R}$, $f(X,Y) = g(XY^{-1})$ is a smooth, right invariant function. Furthermore, if g is a Morse-Bott function with a unique global minimum at e and no further local minima, then $Y \mapsto f(Y,e)$ is Morse-Bott with a global minimum at e and no further local minima.*

Proof: The smoothness of f is obvious. For all $X, Y, Z \in G$ we have $f(XZ, YZ) = g(XY^{-1}) = f(X,Y)$ and thus f is right invariant. The second statement follows directly from $f(Y,e) = g(Y)$ for all $Y \in G$. $\square$

An analogous construction yields a left-invariant cost function. Furthermore, we can obtain a right invariant cost function from any left invariant cost function and vice-versa from the following result.

Proposition 22 *Let G be a Lie group and let $f\colon G \times G \to \mathbb{R}$ be a left invariant function. Then $\tilde{f}\colon G \times G \to \mathbb{R}$ defined by*

$$\tilde{f}(X,Y) = f(X^{-1}, Y^{-1})$$

is a right invariant function. Furthermore, if $Y \mapsto f(Y,e)$ is a Morse-Bott function such that e is the only local minimum and a global minimum, then $Y \mapsto \tilde{f}(Y,e)$ has the same properties.

Proof: The right invariance is checked by a straight forward calculation. The second statement follows from the fact that the map inv: $G \to G$, $X \mapsto X^{-1}$ is a global diffeomorphism. $\square$

5.4 Example: Pose estimation on SE(3)

As an application of the cost function construction above, let us recall Example 2, the special Euclidean group SE(3). The system on SE(3) is given by

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} R(t) & p(t) \end{pmatrix} = \begin{pmatrix} R\Omega & RV \end{pmatrix},$$

where R resp. p are the attitude resp. position of a coordinate frame fixed to a rigid body in 3D-space relative to an inertial frame, Ω denotes the angular velocity and V denotes the linear velocity measured in the body-fixed frame. As mentioned before, the natural cost function (23) on SE(3) is not right invariant, and hence Theorem 15 cannot be applied. However, Proposition 21 yields a construction procedure for a right invariant cost function. For this we need to choose a suitable function g on SE(3). We use

$$g(R,p) = \frac{1}{2}\left(\|R - I\|^2 + \|p\|^2\right).$$

It is straightforward (although tedious) to verify that g is a Morse-Bott function with a unique global minimum at $(I, 0)$ and no further local minima. Proposition 21 yields the right invariant cost function

$$\begin{aligned} f((\hat{R},\hat{p}),(Y,y)) &= g((\hat{R},\hat{p})(Y,y)^{-1}) \\ &= \frac{1}{2}\left(\|\hat{R}-Y\|^2 + \|\hat{p}-\hat{R}Y^\top y\|^2\right). \end{aligned}$$

Note that

$$f((\hat{R},\hat{p}),(Y,y)) = \frac{1}{2}\left(\|\hat{R}^\top - Y^\top\|^2 + \|-\hat{R}^\top\hat{p} - (-Y^\top y)\|^2\right).$$

Hence we can view f also as the result of applying Proposition 22 to the cost function (23). In order to construct a left observer, such that Theorem 15 can be applied, we also need a right invariant Riemannian metric. Note that $T_{(I,0)}R_{(\hat{R},\hat{p})}(\Omega, V) = (\Omega\hat{R}, V + \Omega\hat{p})$ for all $(\hat{R},\hat{p}) \in \mathrm{SE}(3)$ and $(\Omega, V) \in \mathfrak{se}(3)$. Hence we can define a Riemannian metric by

$$\langle(\Omega_1\hat{R}, V_1 + \Omega_1\hat{p}), (\Omega_2\hat{R}, V_2 + \Omega_2\hat{p})\rangle = \mathrm{tr}(\Omega_1^\top\Omega_2) + V_1^\top V_2^\top$$

for all $(\Omega_1\hat{R}, V_1+\Omega_1\hat{p}), (\Omega_2\hat{R}, V_2+\Omega_2\hat{p}) \in T_{(\hat{R},\hat{p})}\,\mathrm{SE}(3)$. Note, that we have used the representation of tangent vectors by right translation of the Lie algebra in this definition. For our filter we have to calculate the gradient with respect to the Riemannian metric. As a first step we derive a closed formula for the differential of f with respect to the first variable.

$$d_1 f((\hat{R},\hat{p}),(Y,y))(\Omega\hat{R}, V + \Omega\hat{p}) = -\,\mathrm{tr}(\Omega\hat{R}Y^T) + \langle\hat{p} - \hat{R}Y^\top y, V\rangle$$

where $\mathbb{P}$ again denotes the orthogonal projection onto the skew-symmetric matrices. Hence we have the following formula for the gradient,

$$\begin{aligned} &\mathrm{grad}_1\, f((\hat{R},\hat{p}),(Y,y)) \\ &= T_{(I,0)}R_{(\hat{R},\hat{p})}(\mathbb{P}(\hat{R}Y^\top), \hat{p} - \hat{R}Y^\top y) \\ &= T_{(I,0)}L_{(\hat{R},\hat{p})}\,\mathrm{Ad}_{(\hat{R},\hat{p})^{-1}}(\mathbb{P}(\hat{R}Y^\top), \hat{p} - \hat{R}Y^\top y). \end{aligned}$$

To obtain a representation of $\mathrm{grad}_1 f$ as left-translated elements of $\mathfrak{se}(3)$ we check that

$$\mathrm{Ad}_{(\hat{R},\hat{p})^{-1}}(\mathbb{P}(\hat{R}Y^\top), \hat{p} - \hat{R}Y^\top y) = (\mathbb{P}(Y^\top\hat{R}), \hat{R}^\top\hat{p} - Y^\top y + \hat{R}^\top\mathbb{P}(\hat{R}Y^\top)\hat{p})$$

Assume now that we measure the angular and linear velocities $(w_\Omega, w_V) = (\Omega, V)$ and the system state $(Y, y) = (R, p)$. Using the construction above we get the left filter

$$\begin{aligned} \dot{\hat{R}} &= \hat{R}w_\Omega - \hat{R}\mathbb{P}(Y^\top\hat{R}) \\ \dot{\hat{p}} &= \hat{R}w_V - (\hat{p} - \hat{R}Y^\top y) - \mathbb{P}(\hat{R}Y^\top)\hat{p}. \end{aligned}$$

By Proposition 21, and since g was a suitable Morse-Bott function, we can apply Theorem 15 and see that the right invariant error for this observer converges to the identity for generic initial values.

To the best of our knowledge, the above observer on SE(3) has not been proposed in the literature before. Previous work by two of the authors [14] yielded an observer on SE(3) with exactly the same structure, but with the last term in the $\hat{p}$-equation, i.e. $-\mathbb{P}(\hat{R}Y^\top)\hat{p}$, replaced by $-\hat{R}Y^\top\mathbb{P}(\hat{R}Y^\top)y$. Moreover, Lyapunov-type stability results were proven in that paper for the observers including or not including that term, respectively.

6 Gradient-like observers

In this section we relax the invariance requirements placed on the cost function we have used to design the innovation term in our gradient observers and consider general cost functions. To provide structure for the observer design we require that the error dynamics exhibit autonomous gradient dynamics. This approach leads to a version of an *internal model principle* for invariant systems on a Lie group. The principle states that observers for invariant systems that exhibit gradient dynamics for the canonical invariant errors will contain an internal model of the observed system and can be decomposed into a synchronous term plus an innovation term analogous to that discussed in Section 4. The innovation term is not itself a gradient term unless the cost function has the invariance properties discussed in Section 5 and we term the resulting observers *gradient-like*. This result sharpens a related structural result in [17] (see also [16]), where less stringent conditions were placed on the error dynamics. Sharpening these requirements allows us to derive an almost global convergence result in a very general context.

Theorem 23 *Let $f\colon G\times G\to\mathbb{R}$ be a smooth cost function. Consider a general left observer for system* (1)*,*

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, Y, w_l, t) \tag{24}$$

with measurements $Y = X$ and $w_l = u$. Then the canonical right invariant error E_r displays gradient dynamics

$$\dot{E}_r = -\operatorname{grad}_1 f(E_r, e), \tag{25}$$

if and only if

$$F_{\hat{X}}(\hat{X}, Y, w_l, t) = \hat{X}w_l - T_{\hat{X}Y^{-1}}R_Y \operatorname{grad}_1 f(\hat{X}Y^{-1}, e).$$

The analogous result holds for right observers for system (2) *with measurements $Y = X$ and $w_r = v$, and the canonical left invariant error E_l. In this case, the dynamics of the left invariant error are gradient dynamics*

$$\dot{E}_l = -\operatorname{grad}_1 f(E_l, e), \tag{26}$$

if and only if

$$F_{\hat{X}}(\hat{X}, Y, w_r, t) = w_r \hat{X} - T_{Y^{-1}\hat{X}} L_Y \operatorname{grad}_1 f(Y^{-1}\hat{X}, e).$$

Proof: Assume that we have an observer (24) with error dynamics (25). We can split $F_{\hat{X}}(\hat{X}, Y, w_l, t) = \hat{X} w_l + \alpha$ into the synchronization term $\hat{X} w_l$ and a remainder term α by defining $\alpha = F_{\hat{X}}(\hat{X}, Y, w_l, t) - \hat{X} w_l$. By Lemma 7 and using the same argument as in the proof of Theorem 8, as well as $w_l = u$, we see that

$$\dot{E}_r = T_{\hat{X}} R_{X^{-1}} \alpha.$$

We see that $\alpha = -T_{\hat{X}Y^{-1}} R_Y \operatorname{grad}_1 f(\hat{X} Y^{-1}, e)$ by using $X = Y$ and (25). In particular, α is an innovation term in the sense of Section 4. If on the other hand $F_{\hat{X}}(\hat{X}, Y, w_l, t)$ has the form as given in the Theorem then a straightforward calculation shows that we get the error dynamics (25). The statement for the error dynamics (26) follows analogously. □

Theorem 23 yields the two gradient-like observers

$$\dot{\hat{X}} = \hat{X} w_l - T_{\hat{X}Y^{-1}} R_Y \operatorname{grad}_1 f(\hat{X} Y^{-1}, e) \tag{27}$$

and

$$\dot{\hat{X}} = w_r \hat{X} - T_{Y^{-1}\hat{X}} L_Y \operatorname{grad}_1 f(Y^{-1}\hat{X}, e). \tag{28}$$

Note that by Lemma 13 these observers coincide with the gradient observers (21) and (22), respectively, if the cost function and the Riemannian metric are right, respectively left, invariant.

Corollary 24 *Assume that $Y \mapsto f(Y, e)$ is a Morse-Bott function with a global minimum at e and no other local minima. Both errors (E_r and E_l) of the left observer* (27) *and both errors of the right observer* (28) *converge to e for generic initial conditions. Furthermore, $f(E_r, e)$ converges monotonically to $f(e, e)$ for the left filter and $f(E_l, e)$ converges monotonically to $f(e, e)$ for the right filter, and this convergence is locally exponential near e in both cases.*

Proof: The proof is entirely analogous to that of Theorem 15. □

Remark 25 *In the case of noisy measurements $w_l = u + \delta$ and $Y = X N_r$, a straightforward calculation shows*

$$\dot{E}_r = \operatorname{Ad}_{\hat{X}} \delta E_r - T_{E_r N_r^{-1}} R_{X^{-1} N_r X} \operatorname{grad}_1 f(E_r N_r^{-1}, e)$$

for the canonical right invariant error of the left gradient-like observer (27)*. The formula for the canonical left invariant error of the right gradient-like observer* (28) *is analogous. It is intuitively clear that these systems will have strong practical stability properties.*

7 Conclusion

This paper provides a coherent theory of nonlinear observer design for systems with left (resp. right) invariant kinematics on a Lie group for which one has a right (resp. left) invariant, non-degenerate, Morse-Bott cost function and full measurements. The key contributions are the observer equations (21) and (22) along with Theorems 14 and 15. The results in Section 5.3 are of practical importance in generating invariant cost functions. Finally, the results presented in Section 6 provide a practical design methodology in the case where a non-invariant cost function is considered. A limitation of the approach described in this paper is the requirement for full measurement of both state and velocity. In work in progress, we are investigating the structure of observers for kinematic systems on Lie groups with partial state measurements.

References

[1] B. Anderson and J. Moore, *Optimal Filtering*. Dover Publications, Inc., 2005.

[2] A. Doucet, N. de Freitas, and N. Gordon, Eds., *Sequential Monte Carlo methods in practice*, ser. Statistics for Engineering and Information Science. Springer, 2001.

[3] S. Salcudean, "A globally convergent angular velocity observer for rigid body motion," *IEEE Transactions on Automatic Control*, vol. 46, no. 12, pp. 1493–1497, 1991.

[4] H. Nijmeijer and T. Fossen, Eds., *New Directions in Nonlinear Observer Design*, ser. Lecture Notes in Control and Information Sciences. Springer, 1999, iSBN: 1852331348.

[5] J. Thienel and R. M. Sanner, "A coupled nonlinear spacecraft attitude controller and observer with an unknow constant gyro bias and gyro noise," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 2011–2015, 2003.

[6] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Complementary filter design on the special orthogonal group SO(3)," in *Proceedings of the 44th Conference on Decision and Control*, Seville, Spain, 2005, pp. 1477–1484.

[7] S. Bonnabel, P. Martin, and P. Rouchon, "A non-linear symmetry-preserving observer for velocity-aided inertial navigation," in *Proceedings of the American Control Conference*, June 2006, pp. 2910–2914.

[8] R. Mahony, T. Hamel, and J.-M. Pflimlin, "Non-linear complementary filters on the special orthogonal group," *IEEE Transactions on Automatic Control*, to appear, accepted for publication October 2007.

[9] A. Tayebi, S. McGilvray, A. Roberts, and M. Moallem, "Attitude estimation and stabilization of a rigid body using low-cost sensors," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 6424–6429.

[10] P. Martin and E. Salaün, "Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 1039–1045.

[11] J. Crassidis, F. Markley, and Y. Cheng, "Nonlinear attitude filtering methods," *Journal of Guidance, Control,and Dynamics*, vol. 30, no. 1, pp. 12–28, January 2007. [Online]. Available: http://www.acsu.buffalo.edu/~cheng3/

[12] B. Vik and T. Fossen, "A nonlinear observer for GPS and INS integration," in *Proceedings of the 40th IEEE Conference on Decision and Control*, 2001, pp. 2956–2961.

[13] H. Rehbinder and B. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 186–199, Feb. 2003.

[14] G. Baldwin, R. Mahony, J. Trumpf, T. Hamel, and T. Cheviron, "Complementary filter design on the Special Euclidean group SE(3)," in *European Control Conference (ECC 2007)*, Kos, Greece, 2007, pp. 3763–3770.

[15] J. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, "Landmark based nonlinear observer for rigid body attitude and position estimation," in *Proceedings of the 46th IEEE Conference on Decision and Control*, 2007, pp. 1033–1038.

[16] S. Bonnabel, P. Martin, and P. Rouchon, "Invariant asymptotic observers," *IEEE Transactions on Automatic Control*, 2008, accepted for publication. [Online]. Available: http://arxiv.org/abs/math.OC/0612193

[17] S. Bonnabel and P. Rouchon, *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems*, ser. Lecture Notes in Control and Information Sciences. Springer-Verlag, 2005, vol. 322, ch. On Invariant Observers, pp. 53–67.

[18] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear observer on lie groups for left-invariant dynamics with right-equivariant output," in *17th IFAC World Congress*, 2008, accepted for publication. [Online]. Available: http://arxiv.org/abs/0707.2286

[19] R. Kalman, "A new approach in linear filtering and prediction problems," *Transactions of the American Society of Mechanical Engineers [ASME], Journal of Basic Engineering*, vol. 82D, pp. 35–45, 1960.

[20] D. Luenberger, "Observing the state of a linear system with observers of low dynamic order," *IEEE Trans. Mil. Electr.*, vol. ME-8, pp. 74–80, 1964.

[21] H. Nijmeijer and I. Mareels, “An observer looks at synchronization,” *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, vol. 44, no. 10, pp. 882–890, 1997. [Online]. Available: http://purl.org/utwente/31085

[22] C. Lageman, R. Mahony, and J. Trumpf, “State observers for invariant dynamics on a lie group,” in *Proceedings of the international conference on the Mathematical Theory of Networks and Systems*, 2008.

[23] N. Bourbaki, *Lie Groups and Lie Algebras. Chapters 1–3*, 2nd ed. Springer, 1989.

[24] U. Helmke and J. Moore, *Optimization and dynamical systems.* Springer, 1994.